# The Twofold Nature of Autonomous Systems: Evidence Combining Stock Market Data with Topological Properties


Simone Mainardi*†, Enrico Gregori* and Luciano Lenzini†
*Institute of Informatics and Telematics (IIT)
Italian National Research Council (CNR), Pisa, Italy
Email: firstname.lastname@iit.cnr.it
†Deptartment of Information Engineering (IET)
University of Pisa, Pisa, Italy
Email: firstname.lastname@iet.unipi.it



***Abstract*—Autonomous Systems (AS) exist and co-exist in two parallel dimensions. In one dimension they are physical networks, whose interconnections are necessary to ensure global Internet reachabilty. In the other dimension, ASes are large well-known companies competing in the same industry. In this paper we bridge together these dimensions by investigating synchronous cross correlations of stock market data and AS-level topological properties. We find that geographically close companies offering similar services are driven by common economic factors. We also provide evidence on the existence and nature of factors governing AS global as well as local topological properties.**


## I. Introduction

The nature of Autonomous Systems (AS) in the Internet is twofold. On the one hand, they are collections of switches and routers intra- and inter-connected via physical links and logical sessions. On the other hand, they are well-established companies that follow complex business strategies to be competitive within the same industry. Although these two natures could seem incommensurable at a first sight, they are actually closely related. Indeed, business strategies entail developing and implementing enterprise policies and plans. The implementation, in the Internet ecosystem, usually consists in operating routers and physical links or in establishing Border Gateway Protocol (BGP) sessions with providers or peers. BGP, the *de facto* standard for Internet traffic exchange, allows companies to finely tune their in- and out-bound traffic according to contracts signed with other companies. Therefore, we argue the existence of a strong mutual coupling between these natures. Strategic management determines changes in the physical links and logical sessions of switches and routers. In turn, the latter connectivity changes affect present and future business strategies. In this paper, we aim at taking a step forward by linking the two natures.

We investigate synchronous cross correlations between stock price variations and AS-level topological properties such as the degree or the clustering coefficient. The AS-level topology is an abstract representation of the economically-driven inter-connections between ASes, which need to cooperate in order to stay on the market. We focus on the AS-level topology since we believe it best captures the dynamics underlying inter-AS economic relationships. Indeed, at this level all the routers and links operated by a single AS are collapsed into one single node and only inter-AS links are retained. These links, corresponding to BGP sessions, are always established according to some kind of economical agreement [1] [2].

Synchronous cross correlations are quantified by means of the Pearson correlation coefficient. A metric space is defined for the investigated stocks and AS-level properties, ensuring that the stronger the correlation, the closer the elements in the space. A hierarchical organization in this space is detected through a clustering procedure able to extract an ultrametric space from it. We emphasize the hierarchical organization by means of a Minimum Spanning Tree (MST), which provides a meaningful topological arrangement of stocks and AS-level properties. We show that this methodology allows to isolate groups which make sense from an economic point of view and provides valuable information on the factors behind the evolution of the Internet ecosystem.

Our contribution can be summarized as follows. We find that groups of companies homogeneous with reference to their service offering – e.g. transit providers – are positively correlated in the stock market. Similarly, we find that even geographically close companies are positively correlated, suggesting the existence of common economic factors driving geographically homogeneous companies. In addition, the topological arrangement obtained through the MST can be used to derive a meaningful taxonomy of the ASes. New evidence on the factors underlying the AS-level topology time evolution is given by combining its properties with stock market data. We highlight the existence of factors, common to all ASes, able to drive the evolution of *global* properties. Other factors, specific for each AS, determine strong correlations between *local* properties. We also show that factors governing AS stock price variations are not the same as those synchronously driving the variations of AS-level properties.

The rest of this paper is structured as follows. In Sect. II we present a brief overview of related work. We describe the methodology in Sect. III and investigated companies in Sect. IV. Results are given in Sect. V and VI. Opportunities and challenges to extend, explore and deepen the ideas proposed are surveyed in Sect. VII, which concludes the paper.

## II. Related Work

Great interest and dedication has been shown so far in the analysis and modeling of the Internet AS-level topology. An intense research activity has begun to emerge after the seminal works [3] and [4]. Analyses and models strongly depend on measurement data provided by projects such as IRL, CAIDA Ark and DIMES. Analyses (e.g. [5] [6] and references therein) rely on measurements to draw meaningful conclusions on the structural properties of the AS-level topology. Models (see [7] for an accurate review an evolutionary comparison) rely on them for validation. Unfortunately, the ability to accurately map the AS-level topology was shown to be fraught with difficulties and dangers [8]. Difficulties are encountered for example when detecting certain kinds of BGP sessions [9] [10] or when inferring the physical devices belonging to each AS [11]. Dangers are due to the "*as-is*" use of available measurement data as good proxies of the real underlying AS-level topology [12].

To overcome these obstacles, researchers designed and deployed novel measurement infrastructures [13] [14], with the aim of providing an increasingly more accurate and detailed view of the topology and its features. However, to the best of our knowledge, stock market data have never been used to augment or refine the knowledge we have of the AS-level topology. Beginning with the pioneering work [15], such data has successfully been used to study and find *topological* arrangements of economically-principled networks and hence we believe it may provide valuable insights also into the AS-level topology structure and evolution. Similarly, although (anti-)correlations have been observed among neighboring AS degrees [16] [17], to the best of our knowledge *cross* correlations between time-evolving AS-level topological properties have never been studied before. We believe they may be relevant for a better understanding of the complex techno-socio-economic factors underlying the Internet. In addition, they may contribute significantly to the design of novel evolutionary or predictive models. Among the economically-principled models we mention the works [7] and [18]. In [7] the authors assume that AS wealth is the result of a multiplicative stochastic process and keep the degree of each AS proportional to its wealth. In the agent-based model proposed in [18] ASes optimize their cost-based fitness function according to provider or peering strategies.

## III. Methodology

We denote with $p_i(d)$ the closure price of stock $i$ at the day $d$. We focus our attention on the logarithmic return of stock $i$, given by $Y_i(d) = \ln p_i(d+1) - \ln p_i(d)$. We quantify the degree of similarity between the synchronous time evolution of a pair of stock price [15] by the Pearson correlation coefficient [19]

$$c_{ij} = \frac{\langle Y_i(d) Y_j(d)\rangle - \langle Y_i(d)\rangle\langle Y_j(d)\rangle}{\sqrt{\langle Y_i^2(d)\rangle - \langle Y_i(d)\rangle^2}\sqrt{\langle Y_j^2(d)\rangle - \langle Y_j(d)\rangle^2}},$$

where the statistical average $\langle \ldots \rangle$ is a temporal average performed on all the trading days of the investigated period. By definition, $c_{ij}$ ranges from $-1$ (completely anti-correlated stocks) to 1 (completely correlated stocks). Stocks are uncorrelated when $c_{ij} = 0$.

Unfortunately, the correlation coefficient does not represent a distance function for any Euclidean space. Therefore, it cannot be used either to build up a hierarchy or to arrange stocks in a topological space. Hence, we use the distance function $d(i,j) = \sqrt{2(1-c_{ij})}$, which defines a metric space by fulfilling the three axioms of an Euclidean distance [20]. We obtain the hierarchical organization of stocks by extracting a ultrametric space [21] from the metric one. Practically, this is achieved via the single-linkage clustering procedure [22], which disposes the stocks on the branches of a unique hierarchical tree. Single-linkage clustering is an ascending, bottom-up aggregation procedure. Initially, each stock is in a different branch and, at each step, the two closest branches are aggregated into one larger branch. Distance between two branches is the minimum distance between any stock of one branch and any stock of the other. We also construct the MST connecting stocks in the metric space, in order to emphasize their hierarchical organization and to arrange them in a topological space. The MST – which alone contains all the information for carrying out single-linkage clustering [23] – gives an alternative way to highlight hierarchies among the investigated stocks.

We are aware that the stock price is an aggregate economical indicator and that other indicators may be able to capture company's situation in more detail – e.g. revenue, sales and investments. Nevertheless, while such indicators are often difficult to obtain, the stock price is publicly available and to some extent condenses in a nutshell several aspects of a company. In the present study we focus on a time span from January 2008 to September 2012. However, we observed that different time spans do not lead to significant changes in the results. We retrieved historical closure prices data from Yahoo![1]. *Monthly* (rather than daily) synchronous cross correlations are considered when combining stock market data with AS-level topological properties – closure prices are averaged on a monthly basis. Due to the incompleteness and the errors affecting AS-level topologies, a daily study of cross correlations would appear to have little meaning. Variations in the AS-level topological properties and their cross correlations are quantified exactly as described above for the stocks.

## IV. Investigated Companies

In this preliminary study we focus our attention on a subset of large, publicly traded companies all over the world owning at least an AS. Typically, all these companies offer a rich portfolio of Internet services. However, each one has a main

[1] http://finance.yahoo.com/

TABLE I
COMPANIES CONSIDERED IN THE STUDY

| Company | Service | H.Q. | Ticker Symbol | Market |
|---|---|---|---|---|
| AT&T | *t1tp* | NA | T | NYSE |
| Verizon | *t1tp* | NA | VZ | NYSE |
| Sprint | *t1tp* | NA | S | NYSE |
| Inteliquent | *t1tp* | NA | IQNT | Nasdaq |
| CenturyLink | *t1tp* | NA | CTL | NYSE |
| Deutsche Telekom | *t1tp* | EU | DTE.DE | XETRA |
| Telecom Italia | *t1tp* | EU | TIT.MI | Milan |
| Telefonica | *t1tp* | EU | TEF.MC | Madrid |
| TeliaSonera | *t1tp* | EU | TLSN.ST | Stockholm |
| NTT | *t1tp* | A | NTT | NYSE |
| Level3 | *t1tp* | NA | LVLT | NYSE |
| TATA Comm. | *t1tp* | A | TATACOMM.NS | Bombay |
| Cogent | *tp* | NA | CCOI | Nasdaq |
| TW Telecom | *tp* | NA | TWTC | Nasdaq |
| Akamai | *cdn* | NA | AKAM | Nasdaq |
| Limelight | *cdn* | NA | LLNW | Nasdaq |
| Rackspace | *cdn* | NA | RAX | NYSE |
| InterNAP | *cdn* | NA | INAP | Nasdaq |
| Equinix | *ixp* | NA | EQIX | Nasdaq |

service which can easily be recognized by looking at its history and activity. Hence, we based our selection on the main service offered and chosen: 14 large IP transit providers; 4 content delivery networks; and 1 internet exchange point. An IP transit provider (*tp*) carries IP traffic, enabling paying customer ASes to reach the whole Internet. If a *tp* has full, free-of-charge Internet reachability, then is termed Tier-1 (*t1tp*). A content delivery network (*cdn*) serves content (e.g. web and multimedia objects) to end-users with high availability and high performance. Content providers pay *cdn*s to better distribute their content among users. An Internet exchange point (*ixp*) is a physical facility that enables Internet companies to directly exchange their traffic, without paying for transit. *ixp*s are mainly used by companies with the aim of reducing their costs by bypassing *t1tp*s. In Tab. I we list selected companies and indicate the geographical location of their headqurters – Europe (EU), North America (NA) or Asia (A). In addition, we report their main service offered, the ticker symbol identifying them in the stock market and the stock market where stocks are traded. Autonomous System Numbers (ASNs) are the following: AT&T (7018), Verizon (701), Sprint (1239), Inteliquent (3257), CenturyLink (209, 3561), Deutsche Telekom (3320), Telecom Italia (6762), Telefonica (12956), TeliaSonera (1299), NTT (2914), Level3 (3356, 3549, 1), TATA Communications (6453), Cogent (174), TW Telecom (4323), Akamai (20940), Limelight (22822), Rackspace (15395), InterNAP (11855) and Equinix ([many]). AS-level topologies are generated using the data available from the Internet Research Lab (IRL) website[2] – outliers in data are discarded using the Chauvenet's criterion [24].

[2]http://irl.cs.ucla.edu/topology/

We chosen 5 AS-level topological properties to analyse the connectivity features of the investigated ASes. The properties are the following:

- *Degree (de):* Is the number of BGP sessions an AS established with other ASes, i.e. with its *neighbors*.
- *Average neighbor degree (knn):* Is the average degree of the neighbors of an AS.
- *Clustering coefficient (cc):* Quantifies how close the neighbors of an AS are to being a clique, i.e. a complete graph.
- *Eigenvector centrality (ei):* Is a measure of the importance of an AS. It assigns relative scores to ASes based on the principle that connections to high-scoring ASes contribute more to the score of the AS in question than equal connections to low-scoring ASes.
- *Coreness (co):* A $k$-core is a maximal subgraph of the AS-level topology in which each AS has at least degree $k$. If an AS belongs to the $k$-core but not to the $(k+1)$-core, then is said to have coreness $k$.

Selected properties are able to capture: direct AS connectivity – *de*; connectivity patterns in the neighborhood of the AS – *knn* and *cc*; and global connectivity features – *ei* and *co*. Indeed, while *de* simply accounts for the number of neighbors an AS has, *knn* and *cc* also tell relevant information on neighbors' BGP connectivity. Specifically, *knn* is an average indicator of the propensity of neighboring ASes to establish BGP sessions with the rest of the network. Similarly, *cc* gives information on the attitude of neighbors in establishing BGP sessions with each other. Since such properties involve only neighboring ASes, we also included global properties *ei* and *co* to quantify the role of each AS in the whole topology. As further discussed in Sect. VI, they both take into account network-wide BGP connectivity features.

## V. AUTONOMOUS SYSTEMS' STOCKS HIERARCHICAL ORGANIZATION

In Fig. 1 we show the MST highlighting the hierarchical organization of investigated stocks – the lower the distance between two companies, the higher and thicker the link connecting them. A first inspection of the MST suggests the existence of two geographically homogeneous groups: Europe (top-left with TEF) and North America (top-to-bottom-right with T). North American companies can be further divided into two smaller-but-stronger subgroups bridged through the link T-EQIX: *t1tp*s and *tp*s with T on the one hand, and *cdn*s with EQIX on the other. Apparently, geography seems not to play a significant role for Asiatic companies TATA and NTT, which connect to the EU and the NA groups, respectively. Actually, its relevance is once again confirmed by looking at their geographical location. Indeed, TATA is headquartered in India, which is much closer to EU rather than NA. Similarly, NTT, which is headquartered in Japan, is much closer to NA rather than EU. The absence of a well-defined Asiatic group may be due to the small number of Asiatic companies investigated. The hierarchical tree of the subdominant ultrametric associated to the MST is shown on the left side of Fig. 2. On the

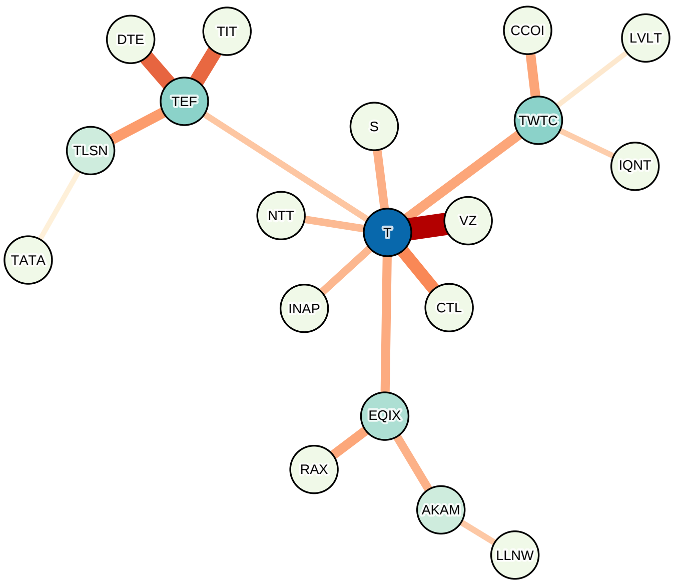


Fig. 1. Autonomous Systems' stocks minimum spanning tree

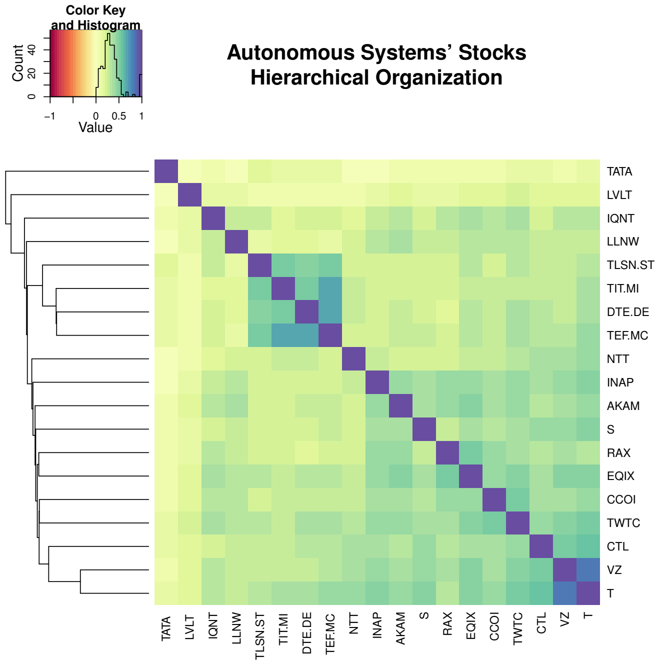


Fig. 2. Autonomous Systems' stocks hierarchical tree

right side we visually represent correlation coefficients using colors by means of an heatmap. Mappings between colors and correlation values are reported in the top-left corner of the figure. In the same corner we also plot an histogram highlighting the correlation coefficients distribution. The lower the correlations a group of stocks has with others, the higher the distance at which the branching occurs in the hierarchical tree. For the sake of example we can consider the branch involving TEF, DTE, TIT and TLSN. It departs early from the rest of the tree and in fact the heatmap highlights very low to no correlations with companies not belonging to the branch. Practically, a branch which departs from the tree at a high distance suggests that the involved companies are subject to *common* economic factors and that these factors do not affect companies outside the branch. To rephrase succinctly, it suggests the existence of economic factors which are *specific* only to companies in the branch. Likewise, when a branching occurs at low distance values – e.g. when VZ and T split into two distinct branches – companies involved are not only subject to *common* economic factors each other. They also have economic factors in *common* with other companies that departed earlier from the same branch – e.g. CTL.

A detailed inspection of the MST and of the branches of the associated hierarchical tree enable to identify two strongly correlated groups in the hierarchy:

- European Tier-1 transit providers (TEF, DTE, TIT, TLSN);
- North American Tier-1 transit providers (T, VZ, CTL).

Such groups correspond to dark diagonal blocks in the heatmap, which in turn map into strongly connected parts of the topological arrangement obtained through the MST. Discovered groups cluster together *t1tp* companies even if (first group) their stocks are traded in different markets. In addition we observe that any company in such groupings is also a telecommunications operator. Once again, we stress on the other striking feature of these groupings, i.e. their geographical homogeneity.

Two less strong groups correspond to:

- North American large, non-Tier-1 transit providers (TWTC, CCOI);
- North American *ixps* and content delivery service providers (EQIX, RAX).

Companies in these smaller groups have less pronounced correlations. Nevertheless, they are topologically close to similar companies in terms of service offerings and headquarters location. Indeed, TWTC and CCOI are close to the north American *t1tp* in the topological arrangement obtained with the MST. This is reasonable if we look at the historical debates about their role as *t1tp*s or simply *tp*s in the Internet. Similarly, EQIX and RAX, are close to north American *cdn*s AKAM and LLNW. EQIX, which is a well-established *ixp*, has tens of datacenters all around the world enabling it to offer also *cdn* services. A similar explanation can be given for INAP, which is connected to *t1tp*s rather than *cdn*s. Indeed, as also pointed out in its services portfolio, it strongly relies on *t1tp*s to distribute contents in the Internet and also offers transit services on its own.

To sum up, the observed groups are meaningful from an economic standpoint since they are composed of companies homogeneous with respect to service offering and geographical location. This empirical evidence suggests the existence of common economic factors driving the synchronous time evolution of geographically homogeneous companies. Additionally, within the same location, companies offering similar services undergo to the same economic factors, which have a service-specific and service-exclusive nature. In contrast, very low to

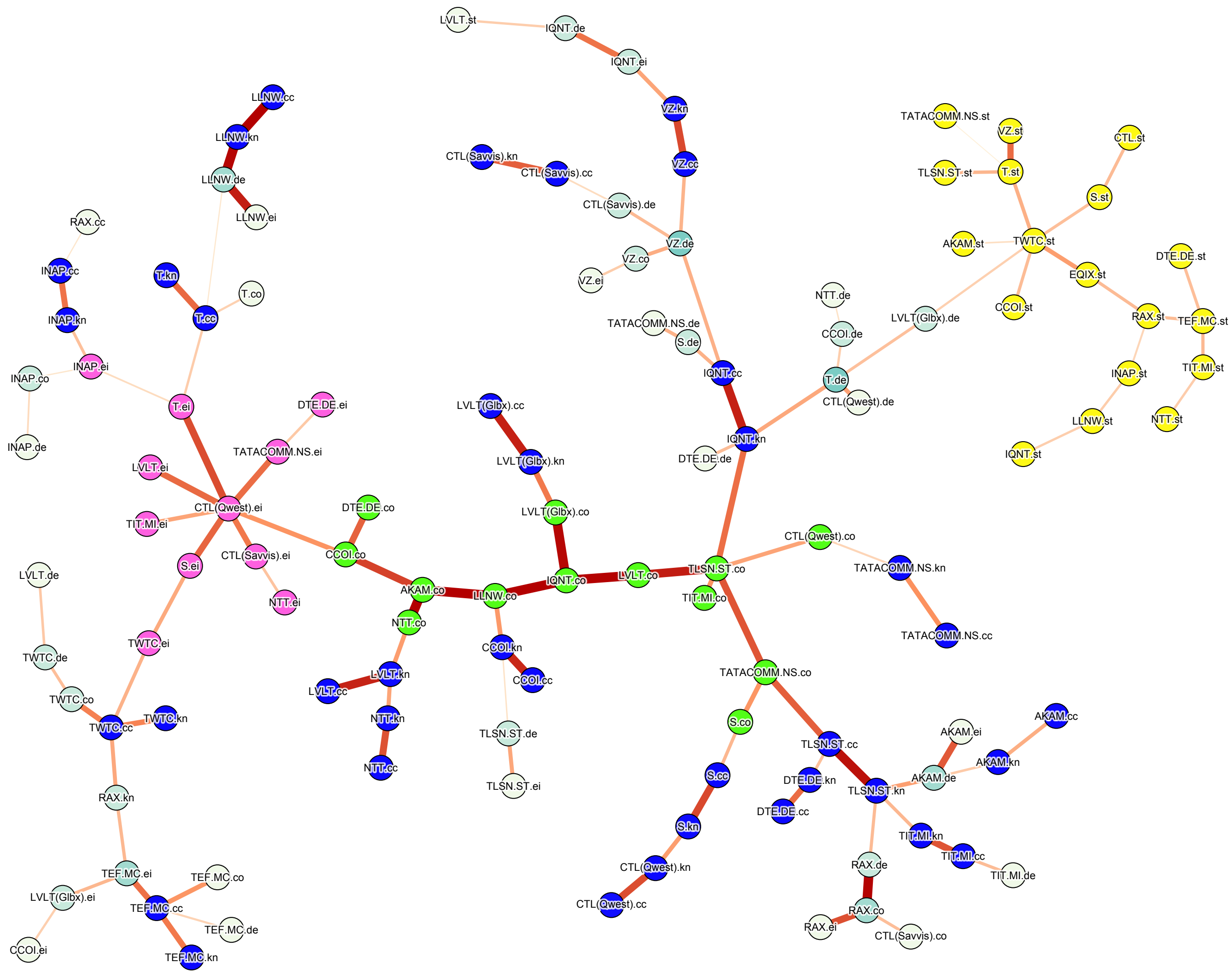


Fig. 3. Autonomous Systems' stocks and topological properties minimum spanning tree

no correlation is found between geographically heterogeneous companies, suggesting that economic factors vary significantly among different countries. Finally, the ability of the MST and the hierarchical tree in isolating homogeneous groups suggests their use in deriving meaningful AS taxonomies.

## VI. Combining AS-level Topological Properties with Stock Prices

In Fig. 3 we show the MST obtained by combining stocks and topological properties. Nodes are labeled with abbreviated, dot-separated company and property names. Space constraints do not allow us to show the hierarchical tree. An inspection of the MST highlights two distinct kinds of groups:

- Large groups, heterogeneous with respect to the company but strongly homogeneous with respect to the property;
- Small, single-company groups of heterogeneous properties (clustering coefficient and average neighbor degree).

Large groups are three and emphasize the presence of synchronous cross correlations among heterogeneous companies with reference to their variations in: stock price (top-right yellow group); coreness (green group in the center); and eigenvector centrality (left star-like purple group). The latter three properties are *global*, uncontrollable and almost completely independent of the single company. They depend on the whole Internet ecosystem. For example, stock prices are influenced by global market trends – their fluctuations do not depend only on the single company. Similarly, coreness and eigenvector centrality depend on the whole AS-level topology, and not on the single AS or on its neighborhood. An AS cannot control its eigenvector centrality (coreness) since it strongly depend on the centrality (coreness) of its neighbors, which in turn depends on the centrality (coreness) of their neighbors, and so on. Therefore, empirical evidence suggest the existence of common, *ecosystem-wide factors* causing simultaneous similar variations of global features among all the Internet companies. For this reason, we argue that the very nature of these factors is embedded in the Internet ecosystem as a whole and not in smaller sub-parts of it. In addition, we observe that global factors underlying stock price variations are not the same as those governing neither the coreness nor the eigenvector centrality – no to very low correlation is measured.

Small, single-company groups (in dark blue) always capture strong correlations between clustering coefficient (*cc*) and average neighbor degree (*knn*) for *each* investigated AS, except for RAX. Strong correlations are highlighted by the thick, dark edges directly connecting *cc* and *knn*. Much lower correlations of these properties are observed among different ASes. Therefore, empirical evidence supports the existence of *company-specific factors* determining per-AS *independent* neighbor selection processes. Indeed, there is a virtually zero chance that over time ASes choose (or are chosen by) either the same neighbors or neighbors with similar BGP connectivity features. If two ASes established BGP sessions with the same neighbors (or with neighbors having similar connectivity features), they would have same values for *cc* and *knn* and maximum positive cross correlation. It follows, therefore, that the aforementioned company-specific factors not only yield independent neighbor selection processes, they also cause each AS to establish BGP sessions with different ASes. In other words, there is a negligible chance that investigated ASes choose (or are chosen by) the same neighbors in the whole AS-level topology.

## VII. Conclusion and Future Directions

In the present paper we investigate synchronous cross correlations between stock market data and AS-level topological properties. We find that groups of companies headquartered in the same location and offering the same services tend to be strongly correlated, suggesting that they are subject, in a statistical way, to the same economic factors. We also discuss on the existence and nature of common factors underlying the evolution in time of AS global and local topological properties.

We believe our novel approach provides valuable insights for example for designing new predictive or evolutionary AS-level models, as well as for validating existing ones. A model may take our results into account in order to design mechanisms able to rewire/grow/shrink the AS-level topology in a way that cross correlations are preserved where necessary. We observed that stock market data cannot be used to infer *synchronous* variations in topological properties. Nevertheless, our study paves the way for a fascinating list of new scientific questions, among which: "What if we consider cross correlations as functions of the time lag?", "Current stock market data can predict future trends in AS-level topological properties (or vice versa)?", "Extending the set of publicly traded Internet companies may lead to new insights into market or AS-level dynamics?", "What if we extend the set of ASes, selecting for example large content providers such as Google or Amazon?", "What if we exploit aggregate indices such as the S&P500 or the gross domestic product of countries?", "May other per-company indicators (e.g. revenue, sales) be used to gain further insights?".


## Acknowledgment

This work has been partially supported by the European Commission within the framework of the CONGAS project FP7-ICT-2011-8-317672.


## References


[1] W. Norton, "Internet service providers and peering," in *Proceedings of NANOG*, vol. 19, 2001.

[2] ——, "The evolution of the us internet peering ecosystem," *Equinix white papers*, 2004.

[3] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM Computer Communication Review*, vol. 29, no. 4, 1999.

[4] R. Albert, H. Jeong, and A. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, 2000.

[5] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, X. Dimitropoulos, A. Vahdat *et al.*, "Lessons from three views of the internet topology," *arXiv preprint cs/0508033*, 2005.

[6] E. Gregori, L. Lenzini, and C. Orsini, "k-dense communities in the internet as-level topology," in *Communication Systems and Networks (COMSNETS), 2011 Third International Conference on*, 2011.

[7] X. Wang and D. Loguinov, "Understanding and modeling the internet topology: economics and evolution perspective," *IEEE/ACM Transactions on Networking (TON)*, vol. 18, no. 1, 2010.

[8] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush, "10 lessons from 10 years of measuring and modeling the internet's autonomous systems," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 9, 2011.

[9] R. Cohen and D. Raz, "The internet dark matter-on the missing links in the as connectivity map," in *Proc. IEEE Infocom*, vol. 6, 2006.

[10] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In search of the elusive ground truth: the internet's as-level connectivity structure," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 36, no. 1, 2008.

[11] K. Chen, D. Choffnes, R. Potharaju, Y. Chen, F. Bustamante, D. Pei, and Y. Zhao, "Where the sidewalk ends: Extending the internet as graph using traceroutes from p2p users," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, 2009.

[12] B. Krishnamurthy, W. Willinger, P. Gill, and M. Arlitt, "A socratic method for validation of measurement-based networking research," *Comput. Commun.*, vol. 34, no. 1, 2011.

[13] A. Faggiani, E. Gregori, L. Lenzini, S. Mainardi, and A. Vecchio, "On the feasibility of measuring the internet through smartphone-based crowdsourcing," in *International Workshop on Wireless Network Measurements (WinMee)*, 2012.

[14] M. Allalouf, E. Kaplan, and Y. Shavitt, "On the feasibility of a large scale distributed testbed for measuring quality of path characteristics in the internet," in *Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops, 2009. TridentCom 2009. 5th International Conference on*, 2009.

[15] R. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 11, no. 1, 1999.

[16] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, "Dynamical and correlation properties of the internet," *Physical review letters*, vol. 87, no. 25, 2001.

[17] J. Park and M. Newman, "Origin of degree correlations in the internet and other networks," *Physical Review E*, vol. 68, no. 2, 2003.

[18] A. Lodhi, A. Dhamdhere, and C. Dovrolis, "Genesis: An agent-based model of interdomain network formation, traffic flow and economics," in *INFOCOM, 2012 Proceedings IEEE*, 2012.

[19] J. Rodgers and W. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, vol. 42, no. 1, 1988.

[20] R. Mantegna and H. Stanley, *Introduction to econophysics: correlations and complexity in finance*. Cambridge University Press, 1999.

[21] R. Rammal, G. Toulouse, and M. Virasoro, "Ultrametricity for physicists," *Reviews of Modern Physics*, vol. 58, no. 3, 1986.

[22] R. Sibson, "Slink: an optimally efficient algorithm for the single-link cluster method," *The Computer Journal*, vol. 16, no. 1, 1973.

[23] J. Gower and G. Ross, "Minimum spanning trees and single linkage cluster analysis," *Applied statistics*, 1969.

[24] V. Barnett and T. Lewis, "Outliers in statistical data," *Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, Chichester: Wiley, 1984, 2nd ed.*, vol. 1, 1984.